

Hardware e Sistemi Operativi

Mirko Cesarini - Dario Pescini
nome.cognome@unimib.it

Università di Milano Bicocca

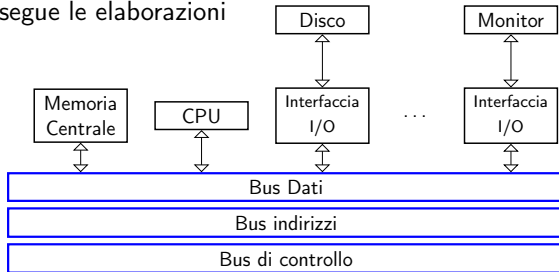
Sistemi Operativi

- (Reminder) *sistema operativo*: si occupa della gestione dell'hardware e delle risorse condivise del computer
- Il sistema operativo fornisce un'*interfaccia* che permette ai programmi applicativi di astrarre dall'hardware
- Prima di analizzare le diverse componenti di un sistema operativo, vediamo alcuni dettagli sull'hardware di un computer

Architettura di Von Neumann

Architetture - I calcolatori esistenti ...

- ... si basano sull'architettura di *Von Neumann*
 - Dati e istruzioni sono ospitati nella *memoria centrale*. Evoluzione rispetto alle architetture precedenti: non si usano memorie separate per dati ed istruzioni, si usa un'unica memoria
 - I trasferimenti di dati e istruzioni avvengono attraverso i bus di sistema
 - I dati possono passare dalla memoria centrale alla CPU, dalla memoria centrale alle periferiche (Monitor, dischi, ...) e viceversa attraverso i bus
 - La CPU esegue le elaborazioni



La CPU è composta da

- ALU. Arithmetic and Logic Unit (unità logico aritmetica). E' la parte della CPU che si occupa delle operazioni numeriche (es., somme, moltiplicazioni, ...) e logiche
- Registri del processore. I registri costituiscono l'interfaccia della CPU con la memoria centrale.
 - Per eseguire un'istruzione (es. l'operazione $2+5$), i dati (2, 5) e l'operatore (+) vengono caricati dalla memoria centrale verso i registri, (questa operazione si chiama *fetching*)
 - L'ALU legge dai registri l'operatore e i dati, esegue l'operazione richiesta salvando il risultato in un registro.
 - Il risultato sarà poi trasferito dal registro alla memoria centrale.
 - Un registro molto importante è il *Program Counter*. Contiene l'indirizzo in memoria dell'istruzione che dovrà essere eseguita subito dopo l'attuale
- CU. La Control Unit (Unità di Controllo, non riportata nello schema precedente) coordina il fetching delle istruzioni, l'esecuzione delle operazioni da parte dell'ALU, la lettura e scrittura dei registri

Clock di sistema

- La maggior parte delle operazioni della CPU possono svolgersi in maniera indipendente l'una dall'altra
 - Es. mentre l'ALU esegue i calcoli di una nuova operazione (A), ...
 - i risultati dell'operazione precedente possono essere trasferiti dai registri del processore alla memoria di massa (B) e ...
 - ... contemporaneamente gli operandi dell'operazione successiva vengono caricati nei registri (C).
 - Il ciclo di operazioni successive può iniziare quando la più lenta tra (A), (B) e (C) termina.
- Il Clock di sistema invia periodicamente un segnale di clock che segnala la fine di un ciclo di operazioni e l'inizio del ciclo successivo. In questo modo le diverse attività vengono coordinate tra loro.
 - La frequenza dei segnali di clock (il numero di segnali inviati in un secondo) misura la velocità con cui una CPU esegue le elaborazioni.
 - La frequenza è misurata in MHz o GHz (Mega Hertz o Giga Hertz, rispettivamente 10^6 o 10^9 cicli al secondo).

Sistema Operativo

Compiti del Sistema Operativo

- Gestire i processi in esecuzione
- Gestire la memoria principale del computer
- Gestire le risorse condivise (periferiche, memorie di archiviazione, ...)
 - Disciplinare l'accesso alle risorse
 - Allocare le risorse ai diversi utilizzatori
- Gestire il File System
- Fornire una user interface (interfaccia utente) per l'interazione tra l'uomo e il computer
- Gestire avviamento e spegnimento del computer
- Gestire l'interconnessione con altri computer (network)
- ...

Gestione dei processi

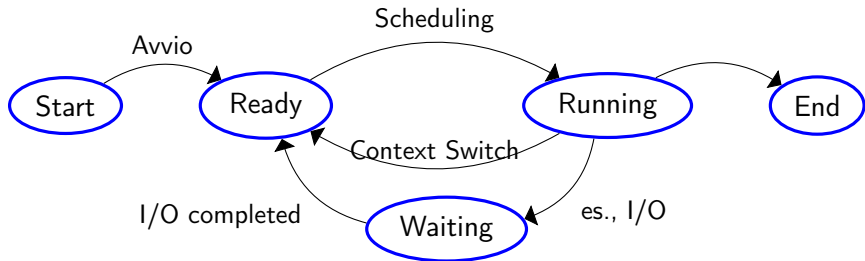
- Un *processo* è un programma in esecuzione.
- Un *processore* è un hardware che esegue i processi
 - Un processo utilizza le risorse fornite dal computer per assolvere i propri compiti
 - Attenzione: un processo non è un programma. Un processo rappresenta l'evoluzione di un programma nel tempo.
- In un elaboratore moderno (pc, smartphone, ...) vengono eseguiti più processi contemporaneamente
- Il sistema operativo è responsabile delle seguenti attività riguardanti la gestione dei processi
 - creazione e terminazione dei processi
 - esecuzione, sospensione e riattivazione dei processi
- Senza una gestione adeguata, la coesistenza delle diverse attività sarebbe difficoltosa

Concorrenza tra processi

- Strategie per eseguire più processi contemporaneamente
 - Multiprogramming
 - Il tempo di elaborazione del processore viene suddiviso tra processi
 - Context switch: un processo in esecuzione viene *messo in pausa* e un altro processo viene eseguito
 - Se i context switch sono frequenti (es diverse centinaia o migliaia in un secondo) sembra che i processi siano eseguiti in parallelo
 - L'esecuzione di più processi su una singolo processore è in realtà una *simulazione*
 - Multiprocessing
 - Il computer dispone di più processori o ...
 - ... di un processore *multicore* (cioè con più *core*). Un processore *multicore* è un processore che al suo interno ha due o più *core*, dove un *core* non è altro che un hardware in grado di eseguire processi). I processori *multicore* non sono altro che un insieme di processori costruiti all'interno dello stesso chip
 - ... o di più processori *multicore*
 - ... o una combinazione delle strategie precedenti

Stati dei processi

- Running. Il processo è in esecuzione sul processore
- Waiting. Il processo è in attesa di qualche evento esterno es., scrittura su file, input da tastiera, pacchetto dati dal network, Il processo non può riprendere l'esecuzione fino all'accadere o al completamento dell'evento atteso
- Ready. Il processo è pronto per essere eseguito. Attende che il sistema operativo lo assegni ad un processore (o ad un core di un processore) per l'esecuzione



Scheduler dei processi

- Lo *scheduler* è il componente del sistema operativo che si occupa di stabilire quale processo (tra quelli disponibili in *ready*) dovrà andare in esecuzione
- Esempio di processi in funzione su un sistema Unix

PID	COMMAND	%CPU	MEM	Context Switch.
626	thunderbird	4.6	404M	3070112+
3157	Google Chrome He	4.1	157M	69637+
483	Terminal	2.3	44M	1067624+
0	kernel_task	1.2	769M	57405148+
141	WindowServer	1.1	163M	5285049+
273	Finder	0.5	464M	2406691+
339	Google Drive	0.3	90M	865105+
337	Dropbox	0.3	151M	2094586+
993	Microsoft Excel	0.0	49M	293471+
94	mysqld	0.0	164M	121+

...

Gestione della memoria principale

- Il sistema operativo si occupa di:
 - tenere traccia delle porzioni di memoria utilizzate e da quali processi
 - allocare e deallocare spazio di memoria in base alle necessità
 - Un processo può richiedere memoria aggiuntiva (es., un processo python che crea una lista di 100'000 elementi
 - Un processo termina e libera la memoria occupata
 - decidere quali processi caricare quando diventa disponibile spazio in memoria
- *Swap*. La memoria occupata da un processo in *waiting* può essere temporaneamente liberata
 - Il contenuto della memoria del processo viene scritta su disco
 - La memoria viene resa disponibile ad altri processi
 - Quando il processo può passare allo stato di *ready* il contenuto precedentemente salvato su disco viene ripristinato in memoria

Memoria virtuale

- *Memoria virtuale*: viene effettuato lo *swap* su disco della memoria centrale occupata da processi in stato di *waiting*. Si può applicare anche a processi in *ready*
 - Vantaggi: è possibile mettere a disposizione dei processi un quantitativo di memoria centrale maggiore di quella fisicamente disponibile
 - Svantaggi: il ricorso alla memoria virtuale rallenta *pesantemente* l'esecuzione di un processo. L'accesso alla memoria di massa è molto più lento dell'accesso alla memoria principale del computer
- La *memoria virtuale* è utile per far fronte a picchi temporanei di richiesta di memoria, se la richiesta si protrae nel tempo... è consigliabile usare altre strategie, es. riducendo il carico di lavoro del computer

Gestione risorse condivise

- Un computer spesso è utilizzato da più utenti
- Ci sono alcune risorse che possono essere utilizzate da un utente per volta (es. la stampante).
 - Il sistema operativo si occupa di gestire le code di stampa (raccogliere le richieste di stampa degli utenti, metterle in coda, mandare in stampa i lavori in coda man mano che la stampante stampa i fogli)
- Nel concedere l'uso di risorse condivise il sistema operativo tiene conto:
 - dei *permessi* attribuiti agli utenti (es., l'accesso alla stampante o a directory di dati può essere concesso solo ad alcuni utenti)
 - delle *quote*. Con il meccanismo delle quote è possibile impostare dei limiti all'uso delle risorse (es. non più di x stampe al mese per utente, non più di y Megabyte di spazio su disco per utente, ...)
 - di *meccanismi di tariffazione*. Se le risorse di calcolo sono acquistate come servizi esterni (es. ore di uso di server fornite in *cloud computing*), la tipologia di contratto stipulata può influire sulle modalità e sulle priorità di allocazione delle risorse

Risorse di calcolo come servizio esterno

- Molte aziende/organizzazioni preferiscono non acquistare computer da adibire a server ma comprare ore di elaborazione fornite in *cloud computing*
- Cloud Computing: le risorse di elaborazione vengono messe a disposizione dell'utilizzatore finale sotto forma di servizio, il quale acquista ore/giorni di elaborazione. I Server utilizzati possono essere anche dall'altra parte del mondo, l'utilizzatore finale li usa in remoto.
- Problema: *make or buy?*
 - *make*: acquisto le risorse per poter produrre internamente un bene o servizio. In questo caso acquisto l'hardware, assumo le persone per poter far funzionare i server, ...
 - *buy*: acquisto il servizio dall'esterno. In questo caso acquisto ore di computazione su server forniti da terze parti

Struttura dei costi

	Acquisto e gestione server	Elaborazione in cloud computing
Costi fissi	Acquisto hardware, impianto condizionamento, costi fissi personale, strutture ridondate, backup, ...	-
Costi periodici	Manutenzioni e riparazioni	-
Costi variabili	Elettricità, costi orari del personale	Costo orario del servizio

- Costi
 - Per un uso saltuario o sporadico, il cloud computing può essere molto conveniente
 - All'aumentare delle ore di utilizzo di un server, occorre valutare se l'alternativa *make* diventa più conveniente dell'alternativa *buy*.

Considerazioni

- Il problema (make or buy) si pone solo per i server, per i personal computer è più conveniente l'acquisto (per ora, in futuro le cose potrebbero cambiare)
- Domanda: secondo voi, perché vi raccontiamo tutto ciò?
 - Vi potrà capitare di dover svolgere delle elaborazioni, ma purtroppo la vostra organizzazione (università, azienda, pubblica amministrazione) potrebbero non disporre di elaboratori con caratteristiche adeguate
 - Spendere migliaia di euro per comprare un server (*make*) potrebbe essere meno conveniente rispetto ad acquistare alcune ore di elaborazione (*buy*).

Amazon EC2

- Caso di studio
 - Amazon Elastic Compute Cloud (Amazon EC2)
 - Maggiori informazioni qua
- Nota bene: non vogliamo fare pubblicità ... tuttavia EC2 è stato uno dei primi esempi di cloud computing disponibili al grande pubblico. Per completezza citiamo qualche altro servizio concorrente:
 - Microsoft Azure
 - Google Cloud Platform
 - IBM Bluemix
 - ... e tanti altri che potete rintracciare usando un motore di ricerca

Amazon EC2

EC2 vi mette a disposizione risorse di calcolo in base a:

- Location geografica dei computer. Questo è un aspetto molto importante se la macchina sarà utilizzata per esempio per realizzare un server web. Es. (lista incompleta)
 - America: Virginia settentrionale, Ohio, Oregon, California settentrionale, Montreal, San Paolo, . . .
 - Europa/Medio Oriente/Africa: Irlanda, Francoforte, Londra
 - Asia Pacifico: Singapore, Tokyo, Sydney, Seul, Mumbai, Pechino
- Tipologia di hardware. Vedi prossime slide
- Tipologia di storage (memoria di massa) utilizzata. Ci sono diverse scelte (es., dischi magnetici, dischi a stato solido, . . .). Non le approfondiremo
- Prestazioni di rete (per la trasmissione dati)

EC2 - Hardware

Nota: nella tabella qua sotto sono riportate come esempi solo alcune delle alternative possibili.

Nomi, caratteristiche e prezzi possono cambiare nel tempo.

Nome	N. CPU	Memoria (Gigabyte)	Prestazioni di rete	Clock (GHz)	Costo orario Dollari/ora
t2.micro	1	1	Basso/mod.	Variab.	\$0.012
t2.small	1	2	Basso/mod.	Variab.	\$0.023
t2.medium	2	4	Basso/mod.	Variab.	\$0.047
t2.large	2	8	Basso/mod.	Variab.	\$0.094
t2.xlarge	4	16	Moderato	Variab.	\$0.188
m4.xlarge	4	16	Alto	2,4	\$0.215
m4.2xlarge	8	32	Alto	2,4	\$0.431
m4.4xlarge	16	64	Alto	2,4	\$0.862
m4.10xlarge	40	160	10 Gigabit	2,4	\$2.155
m4.16xlarge	64	256	20 Gigabit	2,3	\$3.447
x1.32xlarge	128	1.952	20 Gigabit	2,3	\$13.338
...

EC2 - Demo

- Ho appena configurato un server

The screenshot shows the AWS Management Console interface for EC2 instances. At the top, there are buttons for 'Launch Instance', 'Connect', and 'Actions'. Below these is a search bar with the text 'Filter by tags and attributes or search by keyword'. A table lists the instance details:

Name	Instance ID	Instance Type	Availability Zone	Instance State
	i-04f318d1d878763f1	t2.micro	us-west-2b	stopped

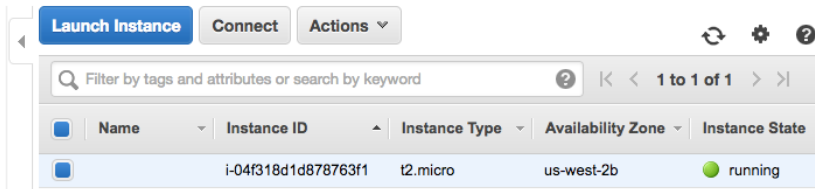
- Avvio il server

The screenshot shows the same AWS Management Console interface, but with the 'Actions' dropdown menu open. The menu options are:

- Connect
- Get Windows Password
- Launch More Like This
- Instance State (highlighted) ▶ Start (highlighted)
- Instance Settings ▶ Stop
- Image ▶ Reboot
- Networking ▶ Terminate
- CloudWatch Monitoring ▶

The instance in the background table is still in the 'stopped' state.

- *Server up and running* (Avviato correttamente)



The screenshot displays the AWS Management Console interface for an EC2 instance. At the top, there are three buttons: 'Launch Instance' (highlighted in blue), 'Connect', and 'Actions' (with a dropdown arrow). To the right of these buttons are three icons: a refresh icon, a settings icon, and a help icon. Below the buttons is a search bar with the placeholder text 'Filter by tags and attributes or search by keyword'. Underneath the search bar is a table with the following columns: 'Name', 'Instance ID', 'Instance Type', 'Availability Zone', and 'Instance State'. The table contains one row with the following data: a blue selection checkbox, an empty 'Name' cell, the Instance ID 'i-04f318d1d878763f1', the Instance Type 't2.micro', the Availability Zone 'us-west-2b', and the Instance State 'running' with a green dot icon.

<input type="checkbox"/>	Name	Instance ID	Instance Type	Availability Zone	Instance State
<input type="checkbox"/>		i-04f318d1d878763f1	t2.micro	us-west-2b	● running

- E ora? Come lo posso utilizzare?

- Posso aprire un terminale remoto ...
 - ...il terminale gira sul mio computer
 - ...ma è collegato al server EC2

```
Last login: Wed Jan 18 18:15:06 2017 from ...
```

```
__| ( __|_ )  
_| ( / Amazon Linux AMI  
__|\__|__|
```

```
https://aws.amazon.com/amazon-linux-ami/...  
[ec2-user@ip-172-31-16-70 ~]$
```

- Posso avviare dei comandi, come se fossi sul mio computer

```
[ec2-user@ip-172-31-16-70 ~]$ python myScript.py
```

- Capite ora l'utilità della riga di comando?

Fine!

- Grazie per l'attenzione!
- Ci sono domande?

That's all Folks!

- Questa è l'ultima lezione del corso!

